

# An extension of Mantel's theorem to random 4-uniform hypergraphs\*

Ran Gu, Xueliang Li, Zhongmei Qin, Yongtang Shi, Kang Yang

Center for Combinatorics and LPMC-TJKLC

Nankai University, Tianjin 300071, P.R. China

Email: guran323@163.com, lxl@nankai.edu.cn, qinzhongmei90@163.com,

shi@nankai.edu.cn, yangkang89@163.com

## Abstract

A sparse version of Mantel's Theorem is that, for sufficiently large  $p$ , with high probability (w.h.p.), every maximum triangle-free subgraph of  $G(n, p)$  is bipartite. DeMarco and Kahn proved this for  $p > K\sqrt{\log n/n}$  for some constant  $K$ , and apart from the value of the constant, this bound is the best possible. Denote by  $T_3$  the 3-uniform hypergraph with vertex set  $\{a, b, c, d, e\}$  and edge set  $\{abc, ade, bde\}$ . Frankl and Füredi showed that the maximum 3-uniform hypergraph on  $n$  vertices containing no copy of  $T_3$  is tripartite for  $n > 3000$ . For some integer  $k$ , let  $G^k(n, p)$  be the random  $k$ -uniform hypergraph. Balogh et al. proved that for  $p > K \log n/n$  for some constant  $K$ , every maximum  $T_3$ -free subhypergraph of  $G^3(n, p)$  w.h.p. is tripartite and it does not hold when  $p = 0.1\sqrt{\log n/n}$ . Denote by  $T_4$  the 4-uniform hypergraph with vertex set  $\{1, 2, 3, 4, 5, 6, 7\}$  and edge set  $\{1234, 1235, 4567\}$ . Pikhurko proved that there is an  $n_0$  such that for all  $n \geq n_0$ , the maximum 4-uniform hypergraph on  $n$  vertices containing no copy of  $T_4$  is 4-partite. In this paper, we extend this type of extremal problem in random 4-uniform hypergraphs. We show that for some constant  $K$  and  $p > K \log n/n$ , w.h.p. every maximum  $T_4$ -free subhypergraph of  $G^4(n, p)$  is 4-partite.

**Keywords:** Mantel's theorem; random hypergraphs; Turán number.

**AMS subject classification 2010:** 05C65, 05C35, 05C80, 05C75.

---

\*Supported by NSFC, the "973" program and PCSIRT.

# 1 Introduction

Mantel's theorem [13] is known as a cornerstone result in extremal combinatorics, which shows that every triangle-free graph on  $n$  vertices has at most  $\lfloor n^2/4 \rfloor$  edges and the unique triangle-free graph that achieves this bound is the complete bipartite graph whose partite sets are as equally-sized as possible. In other words, every maximum (with respect to the number of edges) triangle-free subgraph of  $K_n$  is bipartite.

It is natural to generalize Mantel's theorem to hypergraphs. The *generalized triangle*, denoted by  $T_k$ , is a  $k$ -uniform hypergraph with vertex set  $[2k-1]$  and edges  $\{1, \dots, k\}$ ,  $\{1, 2, \dots, k-1, k+1\}$ , and  $\{k, k+1, \dots, 2k-1\}$ . The *Turán hypergraph*  $T_r(n)$  is the complete  $n$ -vertex  $r$ -uniform  $r$ -partite hypergraph whose partite sets are as equally-sized as possible. In particular, Mantel's Theorem states that the maximum triangle-free graph on  $n$  vertices is  $T_2(n)$ . Frankl and Füredi [11] proved that the maximum 3-uniform hypergraph on  $n$  vertices containing no copy of  $T_3$  is  $T_3(n)$  for  $n > 3000$ . In [15], Pikhurko proved that there exists an  $n_0$  such that for all  $n \geq n_0$ , the maximum 4-uniform hypergraph on  $n$  vertices containing no copy of  $T_4$  is  $T_4(n)$ . Actually, there are also some results on the maximum  $k$ -uniform hypergraphs containing no copy of some other specific hypergraphs, such as paths and cycles [7, 12, 14].

DeMarco and Kahn [10] considered a sparse version of Mantel's Theorem. Let  $G$  be the Erdős-Rényi random graph  $G(n, p)$ . An event occurs *with high probability* (w.h.p.) if the probability of that event approaches 1 as  $n$  tends to infinity. It is interesting to determine for what  $p$  every maximum triangle-free subgraph of  $G(n, p)$  is w.h.p. bipartite. DeMarco and Kahn proved that this holds if  $p > K\sqrt{\log n/n}$  for some constant  $K$ , and apart from the value of the constant this bound is the best possible. Problems of this type were first considered by Babai, Simonovits and Spencer [2]. Brightwell, Panagiotou, and Steger [5] proved the existence of a constant  $c$ , depending only on  $\ell$ , such that whenever  $p \geq n^{-c}$ , w.h.p. every maximum  $K_\ell$ -free subgraph of  $G(n, p)$  is  $(\ell-1)$ -partite.

Recently, Balogh et al. [3] studied an extremal problem of this type in random 3-uniform hypergraphs. For  $n \in \mathbb{Z}$  and  $p \in [0, 1]$ , let  $G^r(n, p)$  be a random  $r$ -uniform hypergraph with  $n$  vertices and each element of  $\binom{[n]}{r}$  occurring as an edge with probability  $p$  independently of each other. Note that in particular,  $G^2(n, p) = G(n, p)$  is the usual graph case. Balogh et al. showed that for  $p > K \log n/n$  for some constant  $K$ , every maximum  $T_3$ -free subhypergraph of  $G^3(n, p)$  w.h.p. is tripartite and it does not hold when  $p = 0.1\sqrt{\log n/n}$ .

In this paper, we extend this type of extremal problem in random 4-uniform hyper-

graphs. Denote by  $T$  the 4-uniform hypergraph  $T_4$  with vertex set  $\{1, 2, 3, 4, 5, 6, 7\}$  and edge set  $\{1234, 1235, 4567\}$ . As a sparse version of Pikhurko's result, we obtain the following theorem.

**Theorem 1.1** *There exists a positive constant  $K$  such that w.h.p. the following is true. If  $G = G^4(n, p)$  is a 4-uniform random hypergraph with  $p > K \log n/n$ , then every maximum  $T$ -free subhypergraph of  $G$  is 4-partite.*

We should point out that the proof technique of Theorem 1.1 is similar but more complicated than that of Balogh et al. used in [3]. The following theorem from [16] states asymptotic general structure, which is much useful to our proof of the main result.

**Theorem 1.2** *For every  $\delta > 0$  there exist positive constants  $K$  and  $\epsilon$  such that if  $p_n \geq K/n$ , then w.h.p. the following holds. Every  $T$ -free subhypergraph of  $G^4(n, p_n)$  with at least  $(3/32 - \epsilon) \binom{n}{4} p_n$  edges admits a partition  $(V_1, V_2, V_3, V_4)$  of  $[n]$  such that all but at most  $\delta n^4 p_n$  edges have one vertex in each  $V_i$ .*

Actually, asymptotic general structure statements can also be concluded from the recent results of [4, 8, 9, 16], et al..

The rest of the paper is organized as follows. In Section 2, we introduce some more notation and prove some standard properties of  $G^4(n, p)$ . The main theorem is proved in Section 3. To simplify the formulas, we shall omit floor and ceiling signs when they are not crucial. Undefined notation and terminology can be found in [6].

## 2 Preliminaries

Let  $G$  denote the 4-uniform random hypergraph  $G^4(n, p)$ . The *size* of a hypergraph  $H$ , denoted by  $|H|$ , is the number of hyperedges it contains. We denote by  $q(G)$  the size of a largest 4-partite subhypergraph of  $G$ . We simply write  $x = (1 \pm \epsilon)y$  when  $(1 - \epsilon)y \leq x \leq (1 + \epsilon)y$ . A vertex set partition  $\Pi = (A_1, A_2, A_3, A_4)$  is *balanced* if  $|A_i| = (1 \pm 10^{-10})n/4$  for all  $i$ . Given a partition  $\Pi = (A_1, A_2, A_3, A_4)$  and a 4-uniform hypergraph  $H$ , we say that an edge  $e$  of  $H$  is *crossing* if  $e \cap A_i$  is non-empty for every  $i$ . We use  $H[\Pi]$  to denote the set of crossing edges of  $H$ .

The *link hypergraph*  $L(v)$  of a vertex  $v$  in  $G$  is the 3-graph with vertex set  $V(G)$  and edge set  $\{xyz : xyzv \in G\}$ . The *crossing link hypergraph*  $L_\Pi(v)$  of a vertex  $v$  is the subhypergraph of  $L(v)$  whose edge set is  $\{xyz : xyzv \text{ is a crossing edge of } G\}$ . The *degree*  $d(v)$  of  $v$  is the size of  $L(v)$ , while the *crossing degree*  $d_\Pi(v)$  of  $v$

is the size of  $L_\Pi(v)$ . The *common link hypergraph*  $L(u, v)$  of two vertices  $u$  and  $v$  is  $L(u) \cap L(v)$  and the *common degree*  $d(u, v)$  is the size of  $L(u, v)$ . The *common crossing link hypergraph*  $L_\Pi(u, v)$  of two vertices  $u$  and  $v$  is  $L_\Pi(u) \cap L_\Pi(v)$  and the *common crossing degree*  $d_\Pi(u, v)$  is the size of  $L_\Pi(u, v)$ . Given three vertices  $u, v$  and  $w$ , their *co-neighborhood*  $N(u, v, w)$  is the set  $\{x : xuvw \in G\}$ , and the *co-degree* of  $u, v$  and  $w$  is the number of vertices in their co-neighborhood. Similarly, given two vertices  $u$  and  $v$ , their co-neighborhood  $N(u, v)$  is  $\{xy : xyuv \in G\}$ ; the co-degree of  $u$  and  $v$  is the number of elements in their co-neighborhood.

Given two disjoint sets  $A$  and  $B$ , we use  $[A, B]$  to denote the set  $\{a \cup b : a \in A, b \in B\}$ . In this paper,  $[A, B]$  usually consists of edges. Given a graph or hypergraph  $G$ , let  $G[A, B]$  denote the set  $G \cap [A, B]$ . We say a vertex partition  $\Pi$  with four classes, which we will call a 4-partition, is *maximum* if  $|G[\Pi]| = q(G)$ . Let  $F$  be a maximum  $T$ -free subhypergraph of  $G$ . Clearly  $q(G) \leq |F|$ . Thus, to prove Theorem 1.1, it is sufficient to show that w.h.p.  $|F| \leq q(G)$ . Moreover, we will prove that if  $F$  is not 4-partite, then w.h.p.  $|F| < q(G)$ .

In the following, some propositions of  $G^4(n, p)$  will be stated. We will use the following Chernoff-type bound to prove those propositions.

**Lemma 2.1** ([1]) *Let  $Y$  be the sum of mutually independent indicator random variables, and let  $\mu = \mathbb{E}[Y]$ , the expectation of  $Y$ . For all  $\epsilon > 0$ ,*

$$Pr[|Y - \mu| > \epsilon\mu] < 2e^{-c_\epsilon\mu},$$

where  $c_\epsilon = \min\{-\ln(e^\epsilon(1+\epsilon)^{-(1+\epsilon)}), \epsilon^2/2\}$ .

In the sequel, we use  $c_\epsilon$  to denote the constant in Lemma 2.1.

**Proposition 2.1** *For any  $0 < \epsilon < 1$ , there exists a constant  $K$  such that if  $p > K \log n/n$ , then w.h.p. the co-degree of any triple of vertices in  $G$  is  $(1 \pm \epsilon)pn$ .*

*Proof.* For each triple of vertices  $u, v, w$ , let  $X_{u,v,w}$  denote the number of vertices  $x \in V \setminus \{u, v, w\}$  such that  $xuvw$  is an edge. Letting  $\mu = \mathbb{E}[X_{u,v,w}]$ , we have  $\mu = p(n-3)$ , and by Lemma 2.1,

$$Pr[|X_{u,v,w} - \mu| > \epsilon\mu] < 2e^{-c_\epsilon(n-3)p} < 2e^{-c_\epsilon \frac{np}{2}}.$$

If  $K > 8/c_\epsilon$ , then  $e^{-c_\epsilon \frac{np}{2}} < n^{-4}$ . By the union bound, it follows that with probability at most  $n^3 n^{-4} = n^{-1}$ , the event  $|X_{u,v,w} - \mu| > \epsilon\mu$  holds for some  $\{u, v, w\}$ . Therefore, w.h.p. there is no such  $\{u, v, w\}$ .  $\square$

**Proposition 2.2** *For any  $0 < \epsilon < 1$ , there exists a constant  $K$  such that if  $p > K \log n/n$ , then w.h.p. the co-degree of any pair of vertices in  $G$  is  $(1 \pm \epsilon) \frac{p}{2} n^2$ .*

*Proof.* For each pair of vertices  $u$  and  $v$ , let  $X_{u,v}$  be the random variable given by the number of pairs of vertices  $x, y \in V \setminus \{u, v\}$  such that  $xuvy$  is an edge. Letting  $\mu = \mathbb{E}[X_{u,v}]$ , we have  $\mu = p \binom{n-2}{2}$ , and then by Lemma 2.1,

$$\Pr[|X_{u,v} - \mu| > \epsilon\mu] < 2e^{-c_\epsilon p \binom{n-2}{2}} < 2e^{-c_\epsilon \frac{n^2 p}{4}}.$$

If  $K > 8/c_\epsilon$ , then  $e^{-c_\epsilon \frac{np}{2}} < n^{-4}$ . By the union bound, it follows that the probability that  $|X_{u,v,w} - \mu| > \epsilon\mu$  for some  $\{u, v, w\}$  is at most  $n^3 2e^{-c_\epsilon \frac{n^2 p}{4}} = o(1)$ .  $\square$

**Proposition 2.3** *For any  $0 < \epsilon < 1$ , there exists a constant  $K$  such that if  $p > K \log n/n$ , then w.h.p. the common degree of any pair of vertices in  $G$  is  $(1 \pm \epsilon) \frac{p^2}{6} n^3$ .*

*Proof.* For two disjoint sets of vertices  $\{u, v\}$  and  $\{a, b, c\}$ , let  $\mathcal{A}_{u,v}^{a,b,c}$  be the event  $\{uabc \in G, vabc \in G\}$ , and let  $X_{u,v}^{a,b,c}$  be the indicator random variable of  $\mathcal{A}_{u,v}^{a,b,c}$ . Then  $d(u, v) = \sum_{abc} \mathcal{A}_{u,v}^{a,b,c}$ . Letting  $\mu = \mathbb{E}[d(u, v)]$ , we have  $\mu = p^2 \binom{n-2}{3}$ , and by Lemma 2.1,

$$\Pr[|d(u, v) - \mu| > \epsilon\mu] < 2e^{-c_\epsilon p \mu} < e^{-c_\epsilon \frac{n^3 p^2}{12}}.$$

By the union bound, it follows that the probability that  $|d(u, v) - \mu| > \epsilon\mu$  for some  $\{u, v\}$  is at most  $n^2 e^{-c_\epsilon \frac{n^3 p^2}{12}} = o(1)$ .  $\square$

**Proposition 2.4** *For any  $0 < \epsilon < 1$ , there exists a constant  $K$  such that if  $p > K \log n/n$ , then w.h.p. for any vertex  $v$  of  $G$ , its degree  $d(v)$  is  $(1 \pm \epsilon) \frac{p}{6} n^3$ .*

*Proof.* For each vertex  $v$ , let  $X_v$  be the random variable given by the number of triples of vertices  $x, y, z \in V \setminus \{v\}$  such that  $vxyz$  is an edge. Letting  $\mu = \mathbb{E}[X_v]$ , we have  $\mu = p \binom{n-1}{3}$ , and by Lemma 2.1,

$$\Pr[|X_v - \mu| > \epsilon\mu] < 2e^{-c_\epsilon p \mu}.$$

By the union bound, it therefore follows that the probability that  $|d(v) - \mu| > \epsilon\mu$  holds for some  $v$  is at most  $n 2e^{-c_\epsilon p \mu} = o(1)$ .  $\square$

**Proposition 2.5** *For any  $0 < \epsilon < 1$ , there exists a constant  $K$  such that if  $p > K \log n/n$ , then w.h.p. for any 4-partition  $\Pi = (A_1, A_2, A_3, A_4)$  with  $|A_2|, |A_3|, |A_4| \geq \frac{n}{80}$ , and any vertex  $v \in A_1$  we have  $d_\Pi(v) = (1 \pm \epsilon)p|A_2||A_3||A_4|$ .*

*Proof.* Since  $|A_2|, |A_3|, |A_4| \geq n/80$ , so  $|A_2||A_3||A_4| \geq n^3/80^3$ . Let  $\mu = \mathbb{E}[d_\Pi(v)] = |A_2||A_3||A_4|p$ . We have

$$\Pr[|d_\Pi(v) - \mu| > \epsilon\mu] < 2e^{-c_\epsilon \frac{n^3 p}{80^3}}.$$

Then by the union bound, the probability that the statement does not hold is bounded by

$$4^n n \Pr[|d_\Pi(v) - \mu| > \epsilon\mu] < e^{3n + \log n - c_\epsilon \frac{n^3 p}{80^3}} = o(1).$$

□

For a vertex  $v$  and a vertex set  $S$ , let  $\mathcal{E}$  be a subset of  $\{vab \in G : a \in S\}$  satisfying that for any  $a \in S$ , there exists  $W \in \mathcal{E}$  such that  $a \in W$ . Let  $Q$  be a subset of  $L(v)$ . Define  $G_{v,\mathcal{E}}[S, Q] = \{xuvw \in G : x \in S, uvw \in Q, \exists W \in \mathcal{E} \text{ s.t. } x \in W, u, w, z \notin W\}$ . Then for any  $xuvw \in G_{v,\mathcal{E}}[S, Q]$  with  $x \in S$  and  $uvw \in Q$ , we can find a  $T = \{vuvw, xuvw, vab\}$ , where  $vab \in \mathcal{E}$ . The condition  $u, w, z \notin W$  in the definition of  $G_{v,\mathcal{E}}[S, Q]$  guarantees that we can find such a  $T$ .

**Proposition 2.6** *For any  $0 < \epsilon, \epsilon_1, \epsilon_2 < 1$ , there exists a constant  $K$  such that if  $p > K \log n/n$ , then w.h.p. for any choice of  $\{v, S, \mathcal{E}, Q\}$  as above with  $|S| = \epsilon_1 n$ ,  $|Q| = \epsilon_2 p n^3$ , we have  $|G_{v,\mathcal{E}}[S, Q]| = (1 \pm \epsilon)p|S||Q|$ .*

*Proof.* Let  $K_{v,\mathcal{E}}[S, Q] = \{xuvw : x \in S, uvw \in Q, \exists W \in \mathcal{E} \text{ s.t. } x \in W, u, w, z \notin W\}$ . Then for  $K_{v,\mathcal{E}}[S, Q]$  and  $\mathbb{E}[G_{v,\mathcal{E}}[S, Q]]$ , we have the relation as follows.

$$\mathbb{E}[G_{v,\mathcal{E}}[S, Q]] = p|K_{v,\mathcal{E}}[S, Q]|.$$

For  $x \in S$ , let  $d_\mathcal{E}(x) = |\{W \in \mathcal{E} : x \in W\}|$  and  $Q_x = \{abc \in Q : vax \in \mathcal{E} \text{ for some vertex } y\}$ . If  $d_\mathcal{E}(x) > 6pn$ , then clearly  $[x, Q] \subseteq K_{v,\mathcal{E}}[S, Q]$ . If  $d_\mathcal{E}(x) \leq 6pn$ , then by Proposition 2.2 we have  $|Q_x| \leq 12pn \times 2^{\frac{pn^2}{2}} = 12p^2 n^3$ . Clearly  $[x, Q \setminus Q_x] \subseteq K_{v,\mathcal{E}}[S, Q]$ . Hence,

$$|[S, Q]| - |K_{v,\mathcal{E}}[S, Q]| \leq \sum_{x \in S, d_\mathcal{E}(x) \leq 6pn} |Q_x| \leq |S| \times 12p^2 n^3 = 12\epsilon_1 p^2 n^4.$$

On the other hand, we have  $|[S, Q]| = |S||Q| = \epsilon_1 \epsilon_2 p n^4$ , so  $|K_{v,\mathcal{E}}[S, Q]| = (1 - o(1))|S||Q|$ . Let  $\mu = \mathbb{E}[G_{v,\mathcal{E}}[S, Q]] = p|K_{v,\mathcal{E}}[S, Q]| = (1 - o(1))p|S||Q|$ . By Lemma 2.1 we have

$$\Pr[|G_{v,\mathcal{E}}[S, Q]| - \mu > \epsilon\mu] < 2e^{-c_\epsilon \mu}.$$

We have at most  $n$  choices for  $v$ ,  $\binom{n}{\epsilon_1 n}$  choices for  $S$ ,  $2^{pn^2 \cdot \epsilon_1 n}$  choices for  $\mathcal{E}$  and  $\binom{\frac{1}{3}pn^3}{\epsilon_2 pn^3}$  choices for  $Q$ . Then by the union bound, the probability that the statement of

Proposition 2.6 does not hold is bounded by

$$n \binom{n}{\epsilon_1 n} 2^{m^2 \epsilon_1 n} \binom{\frac{1}{3} p n^3}{\epsilon_2 p n^3} 2e^{-c_\epsilon \epsilon_1 \epsilon_2 p^2 n^4} = o(1).$$

□

**Proposition 2.7** *For any  $0 < \epsilon < 1$ , there exists a constant  $K$  such that if  $p > K \log n/n$ , then w.h.p. the following holds: if  $F$  is a maximum  $T$ -free subhypergraph of  $G$  and  $\Pi$  is a 4-partition maximizing  $|F[\Pi]|$ , then  $|F| \geq (\frac{3}{32} - \epsilon) \binom{n}{4} p$ , and  $\Pi$  is a balanced partition.*

*Proof.* It suffices to prove it when  $\epsilon$  is small. For a partition  $\Pi = (A_1, A_2, A_3, A_4)$ , it is clear that  $|F| \geq q(G) \geq |G[\Pi]|$ . And Proposition 2.5 implies that w.h.p.  $|G[\Pi]| = (1 \pm \epsilon)p|A_1||A_2||A_3||A_4|$  if  $|A_2|, |A_3|, |A_4| \geq \frac{n}{80}$ . Consider a partition satisfying  $|A_1| = |A_2| = |A_3| = |A_4| = \frac{n}{4}$ , then we have  $|F| \geq (\frac{3}{32} - \epsilon) \binom{n}{4} p$ . Also, Theorem 1.2 implies that if  $\Pi$  maximizes  $|F[\Pi]|$ , then  $|G[\Pi]| \geq |F[\Pi]| \geq (\frac{3}{32} - 2\epsilon) \binom{n}{4} p$ .

If  $\Pi$  is not a balanced partition and  $|A_2|, |A_3|, |A_4| \geq \frac{n}{80}$ , then  $|G[\Pi]| \leq (1 + \epsilon)p|A_1||A_2||A_3||A_4| < (\frac{3}{32} - 2\epsilon) \binom{n}{4} p$ . If  $\Pi$  is not balanced and one of  $|A_1|, |A_2|, |A_3|, |A_4|$  is less than  $\frac{n}{80}$ , then Proposition 2.4 implies that  $|G[\Pi]| < \frac{n}{80}(1 + \epsilon)\frac{1}{6}pn^3 < (\frac{3}{32} - 2\epsilon) \binom{n}{4} p$ . Therefore, if  $\Pi$  maximizes  $|F[\Pi]|$ , then  $\Pi$  is balanced. □

Let  $\alpha = 0.35$ . Given a balanced partition  $\Pi = (A_1, A_2, A_3, A_4)$ , let  $P(\Pi) = \{(u, v) \in \binom{A_1}{2} : d_\Pi(u, v) < \frac{\alpha}{32}p^2n^3\}$ . In other words,  $P(\Pi)$  is the set of pairs of vertices in  $A_1$  that have low common crossing degree.

**Proposition 2.8** *There exists a constant  $K$  such that if  $p > K \log n/n$ , then w.h.p. for every balanced partition  $\Pi$ , every vertex  $v$  and every positive constant  $\xi > 0$ , we have  $d_{P(\Pi)}(v) < \frac{\xi}{p}$ , where  $d_{P(\Pi)}(v)$  denotes the number of elements containing  $v$  in  $P(\Pi)$ .*

*Proof.* Let  $\epsilon = 0.1$ . By Proposition 2.5, we assume that  $d_\Pi(v) \geq (1 - \epsilon)p\frac{n^3}{64}$ , and therefore,  $d_\Pi(u, v) < \frac{2\alpha}{1-\epsilon}d_\Pi(v)p$  for  $(u, v) \in d_{P(\Pi)}(v)$ .

If a vertex  $v$  and a balanced cut  $\Pi$  violate the statement of Proposition 2.8, then there are  $S \subseteq V$  and  $Q = L_\Pi(v)$  with  $|S| := s = \lceil \xi/p \rceil$  and  $|G[S, Q]| \leq \frac{2\alpha}{1-\epsilon}|S||Q|p$ . We have at most  $4^n$  choices of  $\Pi$ ,  $n$  choices of  $v$ ,  $\binom{n}{s}$  choices of  $S$ , so the probability of such a violation is at most

$$4^n n \binom{n}{s} e^{-c \cdot \frac{\xi}{p} \cdot p n^3 \cdot p}$$



for some small constant  $c$ , and therefore is  $o(1)$ .  $\square$

The following lemma plays an important role in the proof of Lemma 3.1.

**Lemma 2.2** *Let  $\beta$  and  $r$  be positive integers. For any  $\epsilon > 0$ , there exists a constant  $K$  such that if  $p > K \log n/n$ ,  $\beta \leq \epsilon n$  and*

$$\binom{n}{\beta} \binom{n^3}{r} \exp(-c_\epsilon \epsilon p n r) = o(1). \quad (1)$$

*Then w.h.p. the following holds: for any set of vertices  $A$  with  $|A| \leq \beta$ , there are at most  $r$  triples  $\{u, v, w\} \in \binom{V(G)}{3}$  such that  $|N(u, v, w) \cap A| > 2\epsilon p n$ .*

*Proof.* Fix a vertex set  $A$  of size  $\beta$ . We shall show that there are at most  $r$  triples  $\{u, v, w\} \in \binom{V(G)}{3}$  such that  $|N(u, v, w) \cap A|$  is large. For each triple  $\{u, v, w\}$ , let  $\mathcal{B}(u, v, w)$  be the event that  $|N(u, v, w) \cap A| > 2\epsilon p n$ . By Lemma 2.1,

$$\Pr[\mathcal{B}(u, v, w)] < e^{-c_\epsilon \epsilon p n},$$

for  $c = c_1$  in Lemma 2.1. If  $\{u, v, w\} \neq \{u', v', w'\}$ , then  $\mathcal{B}(u, v, w)$  and  $\mathcal{B}(u', v', w')$  are independent events. Consequently, the probability that  $\mathcal{B}(u, v, w)$  holds for at least  $r$  triples is at most

$$\binom{n^3}{r} \exp(-c_\epsilon \epsilon p n r).$$

There are  $\binom{n}{\beta}$  choices of  $A$ . Therefore, if Eq. (1) holds, then w.h.p. there are at most  $r$  triples  $\{u, v, w\} \in \binom{V(G)}{3}$  such that  $|N(u, v, w) \cap A| > 2\epsilon p n$ .  $\square$

### 3 Proof of Theorem 1.1

Let  $F$  be a  $T$ -free subhypergraph of  $G$ . We want to show that  $|F| \leq q(G)$ . We first state two key lemmas. The first lemma proves  $|F| \leq q(G)$  with some additional conditions on  $F$ . Define the *shadow graph* of a hypergraph  $H$  is a simple graph with  $xy$  an edge if and only if there exists some edge of  $H$  that contains both  $x$  and  $y$ .

**Lemma 3.1** *Let  $F$  be a  $T$ -free subhypergraph of  $G$  and  $\Pi = (A_1, A_2, A_3, A_4)$  be a balanced partition maximizing  $|F[\Pi]|$ . Denote the shadow graph of  $F$  by  $F_S$ . For  $1 \leq i \leq 4$ , let  $B_i = \{e \in F : |e \cap A_i| \geq 2\}$ . There exist positive constants  $K$  and  $\delta$  such that if  $p > K \log n/n$  and the following conditions hold:*

$$(i) \quad \left| \bigcup_{i=1}^4 B_i \right| \leq \delta p n^4,$$



(ii)  $B_1 \neq \emptyset$ ,

(iii) the subgraph of  $F_S$  induced by  $B_1$  is disjoint from  $P(\Pi)$ ,

then w.h.p.  $|F[\Pi]| + 4|B_1| < |G[\Pi]|$ .

*Remark.* We point out that if Condition (ii) does not hold, i.e.,  $|B_1| = 0$ , then clearly  $|F[\Pi]| + 4|B_1| \leq |G[\Pi]|$ . Therefore, Conditions (i) and (iii) imply that  $|F[\Pi]| + 4|B_1| \leq |G[\Pi]|$ , while Condition (ii) implies the strict inequality.

Let  $F_0$  be a maximum  $T$ -free subhypergraph of  $G$ . By Theorem 1.2 and Proposition 2.7, w.h.p. Condition (i) of Lemma 3.1 holds for every  $\delta > 0$ . Without loss of generality, we may assume that  $|B_1| \geq |B_2|, |B_3|, |B_4|$ . If  $F_0$  is not 4-partite, then Condition (ii) of Lemma 3.1 holds. Moreover, if  $P(\Pi) = \emptyset$ , then Condition (iii) of Lemma 3.1 also holds. Therefore, if  $P(\Pi) = \emptyset$  and  $F_0$  is not 4-partite, then we can apply Lemma 3.1 to  $F_0$  and get that  $|F_0| < q(G)$ , a contradiction. Hence if  $P(\Pi) = \emptyset$  for every balanced partition  $\Pi = (A_1, A_2, A_3, A_4)$ , then the proof would be completed. Thus, we need to consider the case that  $P(\Pi) \neq \emptyset$ . The following lemma tells us that if  $P(\Pi) \neq \emptyset$ , then  $\Pi$  is far from being a maximum 4-partition.

**Lemma 3.2** *There exist positive constants  $K$  and  $\delta$  such that if  $p > K \log n/n$ , the 4-partition  $\Pi$  is balanced, and  $P(\Pi) \neq \emptyset$ , then w.h.p.*

$$q(G) > |G[\Pi]| + |P(\Pi)|\delta n^3 p^2. \quad (2)$$

*Remark.* If  $P(\Pi) = \emptyset$ , then clearly  $q(G) \geq |G[\Pi]| + |P(\Pi)|\delta n^3 p^2$ . Therefore, Lemma 3.2 also implies that  $q(G) \geq |G[\Pi]| + |P(\Pi)|\delta n^3 p^2$  for every balanced 4-partition  $\Pi$ . We point out that the constant  $\delta$  in Lemma 3.1 also satisfies Lemma 3.2, which can be verified in the proof of Lemma 3.2.

Now we use Lemmas 3.1 and 3.2 to prove Theorem 1.1. The proofs of those two lemmas are presented in the next two subsections.

Let  $\tilde{F}$  be a maximum  $T$ -free subhypergraph of  $G$ , so  $|\tilde{F}| \geq q(G)$ . To prove Theorem 1.1, it is sufficient to show that  $|\tilde{F}| \leq q(G)$ . Let  $\Pi = (A_1, A_2, A_3, A_4)$  be a 4-partition maximizing  $|\tilde{F}[\Pi]|$ . From Proposition 2.7, we get that  $\Pi$  is balanced. For  $1 \leq i \leq 4$ , let  $\tilde{B}_i = \{e \in \tilde{F}, |e \cap A_i| \geq 2\}$ . Without loss of generality, we may assume  $|\tilde{B}_1| \geq |\tilde{B}_2|, |\tilde{B}_3|, |\tilde{B}_4|$ . Let  $B(\Pi) = \{e \in G : \exists (u, v) \in P(\Pi) \text{ s.t. } \{u, v\} \subset e\}$  and  $F = \tilde{F} - B(\Pi)$ . Then  $F$  satisfies Condition (iii) of Lemma 3.1 and  $\Pi$  maximizes  $|F[\Pi]|$  as well. By Proposition 2.2, we know that w.h.p.  $|B(\Pi)| \leq |P(\Pi)|pn^2$ . Together with Proposition 2.8, we have that w.h.p.  $|B(\Pi)| \leq O(n^3)$ . Combining with Proposition 2.7, we can apply Theorem 1.2 to  $F$  (let  $\delta$  in Theorem 1.2 be the constant  $\delta$  given

in Lemma 3.1). And since  $\Pi$  maximizes  $|F[\Pi]|$ , we can derive that  $F$  and  $\Pi$  satisfy Condition (i) of Lemma 3.1. For  $1 \leq i \leq 4$ , let  $B_i = \{e \in F, |e \cap A_i| \geq 2\}$ . Then we have:

$$\begin{aligned}
|\tilde{F}| &\leq |\tilde{F}[\Pi]| + 4|\tilde{B}_1| \\
&= |F[\Pi]| + 4|B_1| + 4|\tilde{F} \cap B(\Pi)| \\
&\leq |G[\Pi]| + 4|\tilde{F} \cap B(\Pi)| \\
&\leq |G[\Pi]| + 4|B(\Pi)| \\
&\leq |G[\Pi]| + 4|P(\Pi)|\delta p^2 n^3 \\
&\leq q(G),
\end{aligned}$$

where  $\delta$  is the constant in Lemma 3.1. The second inequality follows from Lemma 3.1 and its remark. Note that the constant  $\delta$  in Lemma 3.1 satisfies Lemma 3.2, the last inequality holds because of Lemma 3.2 and its remark. Hence,  $|\tilde{F}| = q(G)$ . So equalities hold throughout the above inequalities. If  $B_1 \neq \emptyset$ , then by Lemma 3.1,  $|F[\Pi]| + 4|B_1| + 4|\tilde{F} \cap B(\Pi)| < |G[\Pi]| + 4|B(\Pi)|$  and if  $P(\Pi) \neq \emptyset$ , then  $|G[\Pi]| + 4|P(\Pi)|\delta p^2 n^3 < q(G)$ , both are contradictions. So both  $B_1$  and  $P(\Pi)$  are empty sets. It follows that  $\tilde{B}_1$  is an empty set. Since we assume that  $|\tilde{B}_1| \geq |\tilde{B}_2|, |\tilde{B}_3|, |\tilde{B}_4|$ , we have that  $|\tilde{B}_1| = |\tilde{B}_2| = |\tilde{B}_3| = |\tilde{B}_4| = 0$ , which implies that  $\tilde{F}$  is 4-partite.

The proof is thus completed. ■

### 3.1 Proof of Lemma 3.1

Let  $\epsilon_1 = \frac{1}{4200}$ ,  $\epsilon_2 = \frac{1}{7200}$ ,  $\delta = \frac{\epsilon_1^3 \epsilon_2}{320 \times 110 \times 16}$ ,  $\epsilon_3 = \frac{16 \times 80 \delta}{\epsilon_1}$ .

Let  $M$  be the set consisting of crossing edges in  $G \setminus F$ . To prove Lemma 3.1, it suffices to prove that  $4|B_1| < |M|$ , so we assume for contradiction that  $|M| \leq 4|B_1| \leq 4\delta p n^4$ . Our aim is to derive a lower bound for  $|M|$ , which contradicts to our assumption.

For each edge  $W = w_1 w_2 w_3 w_4 \in B_1$ , with  $w_1, w_2 \in A_1$ , from Condition (iii), we have  $w_1 w_2 \notin P(\Pi)$ , which implies there are at least  $\frac{\alpha}{32} p^2 n^3$  choices of  $x \in A_2$ ,  $y \in A_3$ , and  $z \in A_4$  such that  $w_1 x y z$  and  $w_2 x y z$  are both crossing edges of  $G$ . By Proposition 2.1, the co-degree of any three vertices is w.h.p. at most  $2pn$ . Hence, there are at least  $\frac{\alpha}{32} p^2 n^3 - 6pn \geq \frac{1}{80} p^2 n^3$  choices of such triple  $(x, y, z)$  satisfying  $w_3, w_4 \notin \{x, y, z\}$ , and then each of these triples  $(x, y, z)$  together with  $W$  form a copy of  $T = \{w_1 w_2 w_3 w_4, w_1 x y z, w_2 x y z\}$  in  $G$ . Since  $F$  contains no copy of  $T$ , at least one of  $w_1 x y z$  and  $w_2 x y z$  must be in  $M$ .

We can obtain a lower bound for  $|M|$  by counting the number of  $T$  in  $G$  that

contain some edge  $W \in B_1$ , since each such  $T$  must contain at least one edge in  $M$ .

Unfortunately, it is not easy to count the number of  $T$  directly. Actually, we will count copies of  $\hat{T}$  instead of  $T$  in  $G$ , where  $\hat{T}$  is a 4-graph on 5 vertices  $\{w_1, w_2, x, y, z\}$  with two crossing edges  $w_1xyz, w_2xyz$  such that there exists  $W \in B_1$  with  $w_1, w_2 \in W \cap A_1$  and  $x, y, z \notin W$ . It is easy to see that every  $\hat{T}$  yields many copies of  $T$  containing some edge  $W \in B_1$ . From the previous argument, we know that for every pair  $w_1, w_2 \in W \cap A_1$  taken from some edge  $W \in B_1$ , there are at least  $\frac{1}{80}p^2n^3$   $\hat{T}$ 's containing  $w_1, w_2$ .

Let  $L$  be the subgraph of  $F_S$  induced by the vertex set  $A_1$ , let  $C = \{x \in A_1 : d_L(x) \geq \epsilon_1 n\}$  and  $D = A_1 \setminus C$ . Let  $C_1$  be the set of  $x \in C$  such that there are at least  $\epsilon_2 pn^3$  crossing edges in  $F$  containing  $x$ . Set  $C_2 = C \setminus C_1$ .

We first prove four claims, which will be needed in the proof of Lemma 3.1.

**Claim 3.1**  $|C| \leq \epsilon_3 n$ .

*Proof.* Since every vertex in  $C$  has degree at least  $\epsilon_1 n$  in  $L$ , we have  $\epsilon_1 n|C| \leq 2|E(L)|$ .

On the other hand, we show that  $|E(L)| \leq 640\delta n^2$ . For each edge  $wu \in E(L)$ , since  $wu \notin P(\Pi)$ , there are at least  $\frac{1}{80}p^2n^3$  choices of  $x \in A_2, y \in A_3, z \in A_4$ , such that  $\{w, u, x, y, z\}$  is the vertex set of a  $\hat{T}$  in  $G$ . So at least one of  $wxyz$  and  $uxyz$  must be in  $M$ . By Proposition 2.1,  $|M| \cdot 2np \geq |E(L)| \frac{1}{80}p^2n^3$ . Since we assume  $|M| \leq 4|B_1| \leq 4\delta pn^4$ , it follows that  $|E(L)| \leq 640\delta n^2$ . Thus,  $|C| \leq \frac{16 \times 80 \delta n}{\epsilon_1} = \epsilon_3 n$ .  $\square$

The next three claims provide lower bounds for  $|M|$  with different parameters.

**Claim 3.2**  $|M| \geq \frac{\epsilon_1 \epsilon_2}{16 \epsilon_3} pn^3 |C_1|$ .

*Proof.* This inequality is trivial if  $|C_1| = 0$ , so we assume  $|C_1| \geq 1$ . For each  $w \in C_1$ , let  $Q_w = \{(x, y, z) \in A_2 \times A_3 \times A_4 : wxyz \in F\}$ . We have the number of neighbors of  $w$  in  $L$ , denoted by  $|N_L(w)|$ , satisfying  $|N_L(w)| \geq \epsilon_1 n$  and  $|Q_w| \geq \epsilon_2 pn^3$ . We will count the number of copies of  $\hat{T}$  with the vertex set  $\{w, u, x, y, z\}$  in  $G$  such that  $w \in C_1, u \in |N_L(w)|, wxyz \in F$  and  $uxyz \in G$ . By Proposition 2.6 with  $v = w, S \subseteq N_L(w)$  and  $|S| = \epsilon_1 n, \mathcal{E} = \{W \in B_1 : w \in W\}$  and  $Q \subseteq Q_w$  with  $|Q| = \epsilon_2 pn^3$ , there are at least  $\frac{1}{2}|S||Q|p$   $\hat{T}$ 's for each  $w \in C_1$ . Thus, the total number of such copies of  $\hat{T}$  is at least

$$\frac{1}{2} \sum_{w \in C_1} \frac{1}{2} |S||Q|p = \frac{1}{4} |C_1| \epsilon_1 n \cdot \epsilon_2 pn^3 \cdot p = \frac{1}{4} \epsilon_1 \epsilon_2 p^2 n^4 |C_1|. \quad (3)$$

We call an edge  $abcd \in M$  *bad* if  $a \in A_1, b \in A_2, c \in A_3, d \in A_4$  and there are at least  $2\epsilon_3 pn$  vertices  $x \in C_1$  such that  $xbcd \in G$ . Since  $|C_1| \leq |C| \leq \epsilon_3 n$  (Claim 3.1),

from Lemma 2.2 with  $\epsilon = \epsilon_3$ ,  $\beta = \epsilon n$ ,  $r = \frac{\log \log n}{p}$ ,  $A = C_1$ , there are at most  $\frac{\log \log n}{p}$  triples  $(b, c, d) \in A_2 \times A_3 \times A_4$  that are in some bad edge. By Proposition 2.1, each such  $(b, c, d)$  is in at most  $\binom{2pn}{2} \hat{T}$ 's. Thus, the number of copies of  $\hat{T}$  estimated in (3) that contain a non-bad edge from  $M$  is at least

$$\frac{1}{4} \epsilon_1 \epsilon_2 p^2 n^4 |C_1| - \binom{2pn}{2} \frac{\log \log n}{p}.$$

Since  $\binom{2pn}{2} \frac{\log \log n}{p} \leq 2p^2 n^2 \frac{\log \log n}{p} = 2pn^2 \log \log n \leq \frac{1}{8} \epsilon_1 \epsilon_2 p^2 n^4 |C_1|$ , there are at least

$$\frac{1}{4} \epsilon_1 \epsilon_2 p^2 n^4 |C_1| - \frac{1}{8} \epsilon_1 \epsilon_2 p^2 n^4 |C_1| = \frac{1}{8} \epsilon_1 \epsilon_2 p^2 n^4 |C_1|$$

copies of  $\hat{T}$  estimated in (3) that contain a non-bad edge from  $M$ . By the definition of bad edges, every such non-bad edge in  $M$  is in at most  $2\epsilon_3 pn$  such copies of  $\hat{T}$ , so we have

$$|M| \geq \frac{1}{8} \frac{\epsilon_1 \epsilon_2 p^2 n^4 |C_1|}{2\epsilon_3 pn} = \frac{\epsilon_1 \epsilon_2}{16\epsilon_3} p n^3 |C_1|.$$

□

The following claim gives us another lower bound for  $|M|$ , its proof is similar to Claim 3.2, but we need more complicated analysis.

**Claim 3.3** *Let  $L'$  be a subgraph of  $L$  such that  $\Delta(L') \leq \epsilon_1 n$ , then*

$$|M| \geq \frac{pn^2}{320\epsilon_1} |E(L')|.$$

*Proof.* Since for each  $wu \in E(L')$ ,  $wu \notin P(\Pi)$ , there are at least  $\frac{1}{80} p^2 n^3$  choices of  $x \in A_2$ ,  $y \in A_3$ ,  $z \in A_4$ , such that  $\{w, u, x, y, z\}$  is the vertex set of a  $\hat{T}$  in  $G$ . So the total number of such  $\hat{T}$  is at least  $\frac{|E(L')|}{80} p^2 n^3$ . And for each of these copies of  $\hat{T}$ , there must be at least one of  $wxyz$ ,  $uxyz$  in  $M$ .

For an edge  $abc \in M$  with  $x \in V(L')$ , we will count copies of  $\hat{T}$  containing edge  $abc$  in  $G$ . Call an edge  $abc$  bad if there exist at least  $2\epsilon_1 pn$   $y \in N_{L'}(x)$  with  $yabc \in G$ . Combining with Proposition 2.2, there exist at most  $\min\{2pn, d_{L'}(x)\}$  vertices  $y \in N_{L'}(x)$  with  $yabc \in G$ . Denote by  $r_x$  the number of  $(a, b, c)$  such that  $abc$  is bad. Therefore, the number of copies of  $\hat{T}$  that contain a non-bad edge from  $M$  is at least

$$\frac{1}{2} \sum_{x \in V(L')} d_{L'}(x) \frac{p^2 n^3}{80} - \sum_{x \in V(L')} r_x \min\{2pn, d_{L'}(x)\}. \quad (4)$$

We will prove  $\frac{1}{4} d_{L'}(x) \frac{p^2 n^3}{80} \geq r_x \min\{2pn, d_{L'}(x)\}$  for every vertex  $x \in V(L')$ . Depending on the value of  $d_{L'}(x)$ , we have three cases.

**Case 1.**  $d_{L'}(x) > 2pn$ .  $\frac{\log n}{p^2 n^2} \leq d_{L'}(x) \leq \epsilon_1 n$ , We apply Lemma 2.2 with  $\beta = \epsilon_1 n$  and  $r = \frac{\log \log n}{p}$  to obtain that  $r_x \leq \frac{\log \log n}{p}$ .

**Case 2.**  $d_{L'}(x) > 2pn$  and  $\frac{\log n}{p^{k+1} n^{k+1}} \leq d_{L'}(x) \leq \frac{\log n}{p^k n^k}$  for some integer  $k \in [2, \frac{\log n}{\log \log n}]$ . We apply Lemma 2.2 with  $\beta = \frac{\log n}{p^k n^k}$  and  $r = \beta/100$  to obtain that  $r_x \leq \frac{\log n}{100 p^k n^k} \leq \frac{p n d_x}{100}$ .

**Case 3.**  $d_{L'}(x) \leq 2pn$ . We apply Lemma 2.2 with  $\beta = 2pn$  and  $r = \frac{p^2 n^2}{100}$  to obtain that  $r_x \leq \frac{p^2 n^2}{100}$ .

One can easily check that, for each of the above three cases,

$$\frac{1}{4} d_{L'}(x) \frac{p^2 n^3}{80} \geq r_x \min\{2pn, d_{L'}(x)\}.$$

Therefore, the number of copies of  $\hat{T}$  estimated in (4) is at least

$$\frac{1}{4} \sum_{x \in V(L')} d_{L'}(x) \frac{p^2 n^3}{80} = \frac{1}{2} \frac{|E(L')|}{80} p^2 n^3.$$

Bearing in mind that a non-bad edge is in at most  $2\epsilon_1 pn$  copies of  $\hat{T}$  estimated in (4), thus, we have

$$|M| \geq \frac{1}{2} \frac{|E(L')|}{80} p^2 n^3 \cdot \frac{1}{2\epsilon_1 pn} = \frac{|E(L')|}{320\epsilon_1} p n^2.$$

□

**Claim 3.4**  $|M| \geq \frac{pn^3}{130} |C_2|$ .

*Proof.* For every vertex  $x \in C_2$ , the number of edges in  $F[\Pi]$  containing  $x$  is at most  $\epsilon_2 pn^3$ . On the other hand, by Proposition 2.5, w.h.p. the crossing degree of  $x$  in  $G$  is at least  $\frac{pn^3}{65}$ . Thus, there are at least  $\frac{pn^3}{65} - \epsilon_2 pn^3 \geq \frac{pn^3}{130}$  edges of  $M$  incident to  $x$ , so  $|M| \geq \frac{pn^3}{130} |C_2|$ . □

Next we present the proof of Lemma 3.1. First we divide the edges in  $B_1$  into three classes. Set

$$B_1^{(1)} = \{e \in B_1 : |e \cap C| \leq 3, |e \cap C_1| \geq 1\},$$

$$B_1^{(2)} = \{e \in B_1 \setminus B_1^{(1)} : |e \cap C| \leq 3, |e \cap C_2| \geq 1\},$$

$$B_1^{(3)} = \{e \in B_1 \setminus (B_1^{(1)} \cup B_1^{(2)})\}.$$

Considering the cardinality of  $B_1^{(1)}, B_1^{(2)}, B_1^{(3)}$ , we have the following three cases. Recall that we have proved three different lower bounds for  $|M|$ , we will show in any one of the following three cases, a lower bounds for  $|M|$  is larger than  $4|B_1|$ .

**Case 1.**  $3|B_1^{(1)}| \geq |B_1|$ .

For any vertex  $x \in C_1$  and  $y \in V(G) \setminus C$ , the co-degree of  $x$  and  $y$  is at most  $pn^2$  by Proposition 2.2. Since the choices of  $y$  is less than  $n$ , there are at most  $pn^3$  edges in  $B_1^{(1)}$  containing  $x$ . Thus, we have  $|B_1^{(1)}| \leq pn^3|C_1|$ . Bearing in mind that  $|M| \geq \frac{\epsilon_1^2 \epsilon_2}{8\epsilon_3} pn^3|C_1|$ , which is shown in Claim 3.4, we can obtain that  $|M| \geq 13pn^3|C_1| \geq 13|B_1^{(1)}| > 4|B_1|$ .

**Case 2.**  $3|B_1^{(2)}| \geq |B_1|$ .

Every  $x \in C_2$  is in less than  $\epsilon_2 pn^3$  crossing edges of  $F$ . Note that every edge in  $B_1^{(2)}$  has at least one vertex in  $C_2$ . If there exist more than  $\epsilon_2 pn^3$  edges with form  $xyuw$ , where  $y \in A_1$ ,  $u \in A_2$ ,  $w \in A_3$ , then we can increase the number of crossing edges across the partition by moving  $x$  to  $A_4$ . Thus, in  $F$ , there are at most  $\epsilon_2 pn^3$  edges of form  $xyuw$ , where  $y \in A_1$ ,  $u \in A_2$ ,  $w \in A_3$ . Furthermore, we can deduce that the total number of edges with form  $xyuw$ , where  $y \in A_1$ ,  $u \in A_i$ ,  $w \in A_j$ ,  $i \neq j \in \{2, 3, 4\}$ , is at most  $3\epsilon_2 pn^3$ .

For any  $x \in C_2$  and  $y \in N_L(x)$ , we count the edges with form  $xyuw$ , where  $u, w \in A_2$ . Realize that if there exists a vertex  $u \in A_2$ , such that there  $xyuw_1, xyuw_2, \dots, xyuw_{\epsilon_1 pn} \in F$ , where  $u, w_1, \dots, w_{\epsilon_1 pn} \in A_2$ , then for any  $x' \neq x \in C_2$  and  $y' \neq y \in A_1$ ,  $x'y'w_iw_j$  is not in  $F$ , where  $i \neq j$ , otherwise, there exists a  $T$  in  $F$ . Since we want to prove  $|M| > 4|B_1|$ , for every vertex  $u \in A_2$ , we assume the number of edges  $xyuw$ , where  $w \in A_2$  is less than  $\epsilon_1 pn$ . So the number of edges  $xyuw$ , where  $x \in C_2$ ,  $y \in N_L(x)$ ,  $u, w \in A_2$ , is at most  $\frac{1}{2} \frac{n}{4} \epsilon_1 pn \frac{n}{4} |C_2| = \frac{\epsilon_1}{32} n^3 p |C_2|$ . Thus, the total number of edges  $xyuw$ , where  $x \in C_2$ ,  $y \in N_L(x)$ ,  $u, w \in A_i$ ,  $i = 2, 3, 4$ , is at most  $\frac{3\epsilon_1}{32} n^3 p |C_2|$ .

Now we count the edges  $xyzu$ , where  $x, y, z \in C_2$ ,  $u \notin C$ . Similarly, since  $|C_2| \leq |C| \leq \epsilon_3 n$ , for every fixed  $x \in C_2$  there are at most  $\binom{\epsilon_3 n}{2}$  choices of pair  $(y, z)$ . Combining with Proposition 2.1, the number of such edges is at most  $\binom{\epsilon_3 n}{2} 2np |C_2| \leq \epsilon_3^2 n^3 p |C_2|$ . Therefore, we have  $|B_1^{(2)}| \leq 3\epsilon_2 pn^3 + \frac{3\epsilon_1}{32} n^3 p |C_2| + \epsilon_3^2 n^3 p |C_2|$ . So  $|M| \geq \frac{1}{130} n^3 p |C_2| > 13|B_1^{(2)}| > 4|B_1|$ .

**Case 3.**  $3|B_1^{(3)}| \geq |B_1|$ .

Let  $L' = L[C] \cup L[D]$ . For any vertex  $x \in D$ ,  $d_L(x) \leq \epsilon_1 n$ . And for any  $y \in C$ , from Claim 3.1, we have  $d_{L'}(y) \leq |C| \leq \epsilon_3 n < \epsilon_1 n$ . Applying Proposition 2.2, we have  $|B_1^{(3)}| \leq pn^2 |E(L')|$ . Combining with Claim 3.3,  $|M| \geq \frac{|E(L')|}{320\epsilon_1} pn^2 > 13|B_1^{(3)}| > 4|B_1|$ .

We complete the proof of Lemma 3.1. ■

### 3.2 Proof of Lemma 3.2

Let  $\epsilon = 0.1$ ,  $\alpha = 0.35$ ,  $\xi = 0.001$ ,  $\gamma = \frac{1-\epsilon}{64} = 0.146$ ,  $\alpha' = \frac{2\alpha}{1-\epsilon} = \frac{7}{9}$ ,  $\varphi = 0.0001$ . Note that for a balanced partition  $\Pi = (A_1, A_2, A_3, A_4)$ ,  $P(\Pi) = \{(u, v) \in \binom{A_1}{2} : d_\Pi(u, v) < \frac{\alpha}{32}p^2n^3\}$ . By Propositions 2.3 and 2.5, we have  $d(u, v) \leq (1 + \epsilon)\frac{p^2}{6}n^3$  for every pair  $(u, v)$  of vertices, and  $d_\Pi(v) \geq (1 - \epsilon)\frac{p}{64}n^3$  for any vertex  $v$ . So  $d_\Pi(u, v) < \frac{2\alpha}{1-\epsilon}pd_\Pi(v) = \alpha'pd_\Pi(v)$  for every pair  $(u, v)$  in  $P(\Pi)$ . Let  $\mathcal{A}$  be the event that for  $\delta > 0$ , there exists a balanced cut  $\Pi$  such that  $q(G) \leq |G[\Pi]| + |P(\Pi)|\delta n^3p^2$ . It is sufficient to prove  $Pr[\mathcal{A}] = o(1)$  for some  $\delta > 0$ . In fact, we will show that  $Pr[\mathcal{A}] = o(1)$  for  $\delta < \gamma\varphi/2$ . It is well known that every graph contains an induced subgraph with at least half of its edges. If we consider  $P(\Pi)$  as an edge set of a graph, then we can derive that there exists a subset  $R$  of  $P(\Pi)$  with  $|R| \geq \frac{1}{2}|P(\Pi)|$ , such that  $R$  is the edge set of some bipartite graph. By Proposition 2.8, we have  $d_{P(\Pi)}(v) < \frac{\xi}{p}$  for every vertex  $v$ . Therefore, we have

$$d_R(v) < \frac{\xi}{p}. \quad (5)$$

Let  $X, Y$  be disjoint subsets of  $V$ ,  $R$  be the edge set of a spanning subgraph of  $[X, Y]$  satisfying (5), and  $f$  be a function from  $X$  to  $\{k \in \mathbb{N} : k \geq \gamma pn^3\}$ . Denote by  $\mathcal{B}(R, X, Y, f)$  the event that there is a balanced cut  $\Sigma$  of  $G$  such that for every vertex  $x$  in  $X$  satisfying

$$d_\Sigma(x) = f(x), \quad R \subseteq P(\Sigma), \quad (6)$$

and

$$q(G) \leq |G[\Sigma]| + \varphi|R|\gamma n^3p^2.$$

We will show that there exists a constant  $c$  such that

$$Pr[\mathcal{B}(R, X, Y, f)] \leq e^{-c|R|n^3p^2}. \quad (7)$$

Note that if  $\delta < \varphi\gamma/2$ , then the event  $\mathcal{A}$  implies event  $\mathcal{B}(R, X, Y, f)$  holds for some choice of  $(R, X, Y, f)$ . Realize that there are at most  $\binom{n}{t}2^t n^{3t}$  choices of  $(R, X, Y, f)$  with  $|R| = t$ . Hence, by the union bound, if (7) holds we have

$$Pr[\mathcal{A}] \leq \sum_{t>0} \binom{n}{t} 2^t n^{3t} e^{-ctn^3p^2} = o(1).$$

Now we prove (7). We choose all edges of  $G$  according to the following three stages.

- (a). Choose the quadruples of vertices of  $G$  that contain  $x \in X$ .



- (b). Choose the rest of the quadruples of vertices of  $G$  except those belonging to  $\bigcup_{y \in Y} [y, \bigcup_{xy \in R} L(x)]$ .
- (c). Choose the rest of the quadruples of vertices of  $G$ .

Let  $H$  be the subhypergraph of  $G$  consisting of the edges described in (a) and (b), and let  $\Gamma$  be a balanced cut of  $H$  maximizing  $|H[\Gamma]|$  among balanced cuts  $\Sigma$  satisfying (6). For each  $y \in Y$ , set  $M(y) = \bigcup_{xy \in R} L_\Gamma(x)$ . We have

$$q(G) \geq |G[\Gamma]| = |H[\Gamma]| + \sum_{y \in Y} |G[y, M(y)]|. \quad (8)$$

Since  $d_\Gamma(x) = f(x) \geq \gamma p n^3$ , for any two vertices  $x$  and  $x'$ , we have  $d_\Gamma(x, x') \leq d(x, x') \leq (1 + \epsilon) \frac{p^2 n^3}{6} \leq \frac{p d_\Gamma(x)}{\gamma}$ . Also since  $R$  satisfies (5), so for each  $y \in Y$  we have

$$\begin{aligned} |M(y)| &\geq \sum_{xy \in R} \left[ d_\Gamma(x) - \sum_{x \neq x' \in N_R(y)} d_\Gamma(x, x') \right] \\ &\geq \sum_{xy \in R} \left[ d_\Gamma(x) - d_R(y) \max_{x \neq x' \in N_R(y)} d_\Gamma(x, x') \right] \\ &\geq \sum_{xy \in R} \left[ d_\Gamma(x) - \frac{\xi p d_\Gamma(x)}{\gamma} \right] \\ &\geq \left(1 - \frac{\xi}{\gamma}\right) \sum_{xy \in R} f(x). \end{aligned}$$

Let  $\mu = p \sum_{y \in Y} |M(y)|$ , then  $\mu \geq (1 - \frac{\xi}{\gamma}) p \sum_{y \in Y} \sum_{xy \in R} f(x)$ . Using Lemma 2.1, we know that with probability at least  $1 - e^{-c_\epsilon \mu} \geq 1 - e^{-c|R|n^3 p^2}$ , for constant  $c = c_\epsilon(\gamma - \xi)$ , the sum  $\sum_{y \in Y} |G[y, M(y)]|$  in (8) is at least  $(1 - \epsilon)\mu$ .

On the other hand, for any balanced cut  $\Sigma$ , we have  $d_\Sigma(x, y) < \alpha' p d_\Sigma(x)$  for all  $(x, y) \in P(\Sigma)$ . So for any balanced cut  $\Sigma$  satisfying (6), we have

$$\begin{aligned} |G[\Sigma]| &\leq |H[\Sigma]| + \sum_{y \in Y} \sum_{xy \in R} d_\Sigma(x, y) \\ &\leq |H[\Gamma]| + \sum_{y \in Y} \sum_{xy \in R} d_\Sigma(x, y) \\ &\leq |H[\Gamma]| + \sum_{y \in Y} \sum_{xy \in R} \alpha' p d_\Sigma(x) \\ &\leq |H[\Gamma]| + \alpha' p \sum_{y \in Y} \sum_{xy \in R} f(x). \end{aligned}$$

Recall that, with probability at least  $1 - e^{-c|R|n^3p^2}$ , the sum  $\sum_{y \in Y} |G[y, M(y)]|$  in (8) is at least  $(1 - \epsilon)\mu$ . We have

$$q(G) - |G[\Sigma]| \geq [(1 - \epsilon)(1 - \frac{\xi}{\gamma}) - \alpha']p \sum_{y \in Y} \sum_{xy \in R} f(x) > \varphi|R|\gamma n^3 p^2$$

holds with probability at least  $1 - e^{-c|R|n^3p^2}$ , which proves (7). It is easy to check that the constant  $\delta$  in Lemma 3.1 satisfies  $\delta < \gamma\varphi/2$ , so we can just let  $\delta$  be the constant  $\delta$  in Lemma 3.1. ■

## References

- [1] N. Alon, J. Spencer, *The Probabilistic Method*, Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., third edition, 2008.
- [2] L. Babai, M. Simonovits, J. Spencer, Extremal subgraphs of random graphs, *J. Graph Theory* **14(5)** (1990), 599–622.
- [3] J. Balogh, J. Butterfield, P. Hu, J. Lenz, Mantel’s theorem of random hypergraphs, arXiv:1310.1501.
- [4] J. Balogh, R. Morris, W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.*, in press.
- [5] G. Brightwell, K. Panagiotou, A. Steger, On extremal subgraphs of random graphs, in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 477–485, New York, 2007. ACM.
- [6] J.A. Bondy, U.S.R. Murty, *Graph Theory*, Springer, New York, 2008.
- [7] N. Bushaw, N. Kettle, Turán numbers for forests of paths in hypergraphs, *SIAM J. Discrete Math.* **28(2)** (2014), 711–721.
- [8] D. Conlon, J. Fox, Y. Zhao, Extremal results in sparse pseudorandom graphs, *Adv. Math.* **256** (2014), 206–290.
- [9] D. Conlon, W.T. Gowers. Combinatorial theorems in sparse random sets, arXiv:1011.4310.

- [10] B. DeMarco, J. Kahn, Mantel's theorem for random graphs, *Random Struct. Algor.*, doi: 10.1002/rsa.20535.
- [11] P. Frankl, Z. Füredi, A new generalization of the Erdős-Ko-Rado theorem, *Combinatorica* **3(3–4)** (1983), 341–349.
- [12] Z. Füredi, T. Jiang, T. Hypergraph Turán numbers of linear cycles, *J. Combin. Theory Ser. A* **123(1)** (2014), 252–270.
- [13] W. Mantel, Problem 28, in *Wiskundige Opgaven*, Volume 10, pages 60–61, 1907.
- [14] D. Mubayi, J. Verstraëte, Minimal paths and cycles in set systems, *European J. Combin.* **28** (2007), 1681–1693.
- [15] O. Pikhurko, An exact Turán result for the generalized triangle, *Combinatorica* **28** (2008), 187–208.
- [16] W. Samotij, Stability results for random discrete structures, *Random Struct. Algor.* **44** (2014), 269–289.